UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION VIII
999 18th STREET - SUITE 500
DENVER, COLORADO 80202-2466

Copy 1

SEP 20 1993

Ref: 8HWM-FF

Charles Scharmann
Office of the Program Manager
ATTN: AMXRM-E
Rocky Mountain Arsenal
Commerce City, Colorado 80022-2180

Re: Rocky Mountain Arsenal (RMA) Integrated Endangerment Assessment (IEA)

Dear Mr. Scharmann:

We have reviewed the Proposed Final Integrated Endangerment Assessment. As provided by the Federal Facility Agreement (FFA), Section 30, we invoke or continue dispute on the issues outlined in the enclosure.

Pursuant to Section 30.2 of the FFA, which requires identification of any activities that the Organization believes are affected by the dispute, we have identified the revision of those sections of the DAA that address biota issues (such as the sections on surficial soils and lakes/sediments) as the primary activity affected by this dispute.

In addition to the enclosed dispute items, we are enclosing a short description of items that need to be addressed prior to finalizing the document. This includes for example statements that appear to be incorrect, and errors in tables or figures.

Our contact on this matter is Ms. Vera Moritz at 294-7517.

Sincerely,

Connally E. Mears

Connally E. Mears
EPA Coordinator for RMA Cleanup

Enclosure

cc: Glenn Tucker, ATSDR
Roberta Ober, PMRMA
CPT. Jonathan Potter, DA
Bradley Bridgewater, Department of Justice

Ronel Finley, USFWS
Jeff Edson, CDH
Vicky Peters, Colorado AG's Office
William Adcock, Shell Oil Company

All with enclosure

ISSUE #1: BMF CALIBRATION (DISPUTE RAISED TO SAPC)

The Army's calibration of the ERC model and consequently calculation of the field biomagnification factor (BMF) is a matter that is currently under dispute and due to be considered by the Steering and Policy Committee (SAPC) on September 22, 1993. The BMF is the single most important factor in determining biota criteria. EPA raised dispute on this issue on June 16, 1993. At that time, EPA Region VIII solicited an independent review of the matter by its Office of Research and Development (ORD). The ORD review, which sets forth the Agency's position on this matter, addressed both the BMF calibration/calculation and the estimation of soil concentrations (see 1.2., below).

1.1. CALIBRATION/BMF_{obs}.

Discussions on this issue are found throughout the Proposed Final IEA (IEA), including, at a minimum, the following locations:

Page ES-7, Section 4.0, paragraph 3
Page ES-9, Second full paragraph
Page ES-10, top paragraph, last sentence
Page 1-3, Section 1.1, third bullet
Page 4-1, Section 4.0, second paragraph, summary of section 4.5
Page 4-2, top three lines
Page 4-5, Section 4.1.2.2, first paragraph
Page 4-8, Section 4.3.1, first paragraph
Page 4-10, Section 4.3.2, first paragraph, last 2 sentences
Page 4-12, Section 4.4.2, second paragraph, third sentence
Pages 4-13, and 4-14, Section 4.5
Pages 4-14 and 4-15, Section 4.6
Page 4-16, Section 4.6.2, second sentence
Page 4-17, Section 4.6.2.1, equation 17
Page 4-30, first paragraph
Page 5-10, top paragraph
Appendix C.1, Section C.1.6
Appendix C.1, Section C.1.7
Appendix C.1, Section C.1.8
Page C.2-20, top paragraph, last sentence
Page C.2-23, first full paragraph
Appendix D.1.3

1.2. CALCULATION OF ESC

On September 1, 1993, EPA Region VIII transmitted to the Army the document titled Position Paper on the Use of Least Squares Regression to Calibrate BMFs at the Rocky Mountain Arsenal, prepared by ORD. This position paper sets forth the Agency's position on the dispute. In the context of the ongoing dispute, we highlight recommendations 3(b) and 3(c) from this position paper:

3.b. Investigate the best use of any spatial structure of the data to develop appropriate interpolation distances for ESC. A statistician with expertise in geostatistical analysis (e.g., kriging) should perform this analysis.

3.c. Develop uncertainty estimates for ESC. Estimates should include uncertainty in averaging and interpolation distances, uncertainties in replacing nondetected concentrations, and uncertainties in averaging algorithms.

Pending final resolution of this ongoing dispute by SAPC, EPA's position is that ESC should be recalculated in accordance with these recommendations, field BMFs should be recalculated accordingly, and the uncertainty estimates for ESC should be incorporated into a more quantitative discussion of uncertainty in risk characterization. These recalculations should be integrated with the other procedures recommended by ORD.

Discussion

ESC is the abbreviation for a quantity originally known as the "Estimated Soil Concentration", but now referred to by the Army as "Exposure Soil Concentration". A precise definition of ESC could not be located in the IEA. Section 4.3.1 states that "For soils, average exposure area concentrations were generally estimated as the arithmetic mean (i.e., average) contaminant concentrations in the 0- to 1-foot soil-depth interval within a receptor's exposure range." (p. 4-9). Section C.1.4 states that "Exposure concentrations are the contaminant concentrations in source media (i.e., soils, sediments, and water) that are bioavailable and accessible, i.e., exposure concentrations are the contaminant concentrations to which organisms are exposed." (p. C.1-4). Section C.1.4.1 states that "Soil exposure concentrations were calculated based on an average area-wide concentration (i.e., an arithmetic mean concentration), an area being defined as an organism's estimated foraging or exposure area." (p. C.1-5).

The variable ESC is used in several ways. First, ESCs for locations where biota samples were collected are used in the calculation of BMF_{obs} (Sections 4.5 and C.1.6). Second, ESCs are used to "predict" tissue concentrations and, hence, to calibrate the food web model, leading to the calculation of BMF_{model} (Sections 4.5 and C.1.6). Third, ESCs calculated for a network of grid points across the Arsenal are used to generate tissue concentration maps and to serve as a "reality check" on the calibrated BMFs (Section E.12.7.6). ESC, therefore, is a key variable in the Army's procedures for estimating exposure, estimating risks, and assessing uncertainties.

EPA has expressed objections to the Army's procedures for calculating and using ESC. For example, in Attachment 3 to EPA's dispute letter of June 16, 1993, EPA expressed the opinion that "The values of ESC .... used by the Army are subject to large errors and cannot be used reliably for calculating either the mean BMF or its statistical distribution." Prior to the issuance of the present report, the Army had not presented a clear statement of the methods used for calculating ESC. Periodic presentations at IEA Subcommittee meetings had indicated that the Army's procedures had changed subsequent to the draft IEA, but these presentations were not sufficiently complete to allow for full review.

The IEA presents for the first time a reasonably complete description of the way in which the Army calculates and uses ESC (although there are still some errors or inconsistencies, as pointed out below). After review of the relevant sections of the IEA, EPA objects to several features of the calculation and use of ESC, and of the discussion of uncertainties in ESC presented in Appendix E.

Our concerns address discussions in Sections 4.3, 4.5, 4.6, C.1.3, C.1.4, C.1.5, C.1.6, C.1.7, C.1.8, E.11, and E.12, at a minimum.

1. Replacement of BCRLs. The soil database contained a large number of samples with contaminant levels reported as "below certified reporting limit" (BCRL). The estimation of these values impact the calculation of the BMF. Section C.1.4.1.1 describes the way in which the Army derived replacement values for these soil samples reported as BCRL. For each BCRL, an initial estimate is derived as the weighted average of up to six "hits" (i.e., values above the BCRL) found within a specified search radius. The search radius varied from 400 to 1,200 feet horizontally and 5 feet vertically; weights apparently varied inversely as the square of the distance. Although the IEA does not state how the six "hits" were selected in cases where more than six occur within the search radius, EPA assumes that the Army's algorithm would select the six closest "hits". After the initial estimates were calculated, the procedure was repeated iteratively (BCRL values were not considered in the first estimate, but replacement values would be included in the second and subsequent iterations). If no hits were found within the search radius of a given BCRL, then no replacement was made for BCRL [this statement corrects an apparent typographical error on p. C.1-6]. Despite this procedure, for the modeled COCs, between 23 and 45 percent of the BCRL values did not receive a replacement value (Table C.1-1 on p. C.1-6; please note that there is another Table C.1-1 elsewhere in this section). Although not clearly stated in the IEA, these non-replaced BCRL values apparently were excluded from the next step in the calculation.

Please note the following concerns: (i) It is not clear why this procedure is used to replace BCRL values in soil, when the "robust" method is selected for BCRL values in tissue. This probably means that the TC and ESC distributions are not comparable at the lower tails (BCRL values for soil are replaced by arithmetic averaging, whereas BCRL values for tissue are replaced by fitting to lognormal distributions). (ii) Because the distribution of soil contamination is highly skewed, it is likely that arithmetic averaging of up to six "hits" will frequently lead to initial estimates higher than the CRL. In this case, the Army's procedure would replace the BCRL with the CRL. It appears that this occurred frequently, leading to upward bias in the replaced values. (iii) No account is taken of the spatial structure of the data. Since no account is taken of clustering, the averaging procedure will be biased towards areas with higher sampling density. The Army acknowledges elsewhere in the IEA that the sampling density for soils was much higher in areas that were more contaminated, and that many of the tissue samples were collected in areas of steep gradients in soil contamination (Section E.12.4.2.3). In these circumstances, averaging of the six closest "hits" would lead to systematic upward bias. (iv) The search radii are too large. Even 400 feet is much larger than the spatial scale of variability in soil contamination levels in many areas. (v) The iterative procedure means that surrounding BCRL values are not included in the replacement process unless and until they themselves have been replaced. Although the consequences of this are not clear, this probably also leads to repeated upward adjustments and consequent upward bias. (vi) The exclusion of non-replaced BCRL values from the remainder of the calculation means that soil data from areas with low contamination levels and/or low sampling densities are systematically excluded from the calculation. However, BCRL tissue values from these areas are not so excluded, resulting in bias in estimates of BMF. Since the proportion of non-replaced values is large for all the COCs, this bias may be large. For all these reasons, it appears that the Army's procedure results in systematic upward bias in the replacement values for soils, and hence in the lower tail of the ESC distribution. This would cause an underestimation of BMF.

2. Interpolation onto the RMA-wide Grid. Section C.1.4.1.2 describes the way in which the Army estimated soil concentrations at each point on a 100- by 100-foot grid. The concentration at each grid point is estimated by calculating a weighted average of all soil concentrations (including replaced BCRLs, but not including non-replaced BCRLs) within a specified search radius. The search radius varied from 200 to 1,200 feet, depending on location; weights were proportional to the inverse square of distance. [The description of the procedure is ambiguous, because this section states "Concentrations in each block were estimated for each contaminant again using an inverse squared-distance algorithm", whereas other parts of the IEA indicate that

the concentrations were estimated only for grid points. The last sentence in this paragraph states "Exposure area concentration estimations within each site group used data from the top 1 ft of borings only, and were constrained to borings located within a particular block." [Emphases added.] The last phrase is incomprehensible, because "block" is defined in this paragraph as the 100' x 100' area bounded by grid points, whereas the search radii were stated to be much larger than 100'. Based on statements in other parts of the IEA, EPA assumes that estimates are actually calculated for each grid point (not for each block) and that the averages are carried out over the area bounded by the search radius (not limited to a block).

Please note the following concerns: (i) The description of the procedure is ambiguous and possibly incorrect, as discussed in the previous paragraph. (ii) The calculation takes no account of the spatial structure of the data. Specifically, it takes no account of clustering of the data, nor of the known tendency for sampling density to be higher in areas where contamination levels are higher (see comment (iii) in the previous section). (iii) The search radii are too large (see comment (iv) in the previous section). (iv) There is no limitation on the number of points within the search radius that may be included in the calculation. Hence, in areas of high sampling density, some or many non-contiguous points may be included in the calculation. For these reasons and those stated in the previous section, it appears that the Army's procedure will yield positive bias in estimates of ESC and too much smoothing (low estimates of variance).

3. Averaging within the exposure range. Section C.1.4.1.3 describes the way in which the Army averaged estimated soil concentrations within the exposure range of each sampled organism to yield a point estimate of ESC. The key sentence in the report appears to be incomplete or in error: "The exposure soil concentration (ESC) for each sample individual was calculated as the average of all grid points located with the center of the species-specific exposure zone defined as the location where the sample was collected." (page C.1-9). EPA assumes that this means that a circle was drawn with center at the point of collection and area equal to the species-specific exposure range (Section C.2.5), and that ESC is calculated as the arithmetic mean of the estimated soil concentration at each grid point located within the circle, regardless of the distance of each grid point from the point of collection.

Please note the following concerns: (i) The description of the procedure is incomplete and possibly in error, as stated in the previous paragraph. (ii) It is not clear why the Army calculates ESC indirectly, using the estimated soil concentration at grid points as intermediate steps in the calculation, instead of estimating the soil concentration at the point of collection directly. The indirect calculation may lead to substantial errors, especially in the areas of steep gradients of soil contamination where many of the biota samples were collected (see Section E.12.4.2.3). (iii) The Army does not state what was done for organisms with very small home ranges (such as terrestrial plants, insects and earthworms). These organisms have exposure ranges with radii much less than the 100-foot spacing between grid points. Hence, in many cases, there would be no grid points located within the exposure range circle, so that the Army's stated procedure would give indeterminate results.

4. Inconsistencies in Estimates of ESC. Table 1 in the ORD position paper dated August 18, 1993, demonstrated major discrepancies between estimates of ESC provided to ORD on two different dates. Preliminary estimates supplied to ORD in March, 1993, were much lower than values supplied in August, 1993. At the very least, this indicates that the values of ESC are extremely sensitive to the method used to calculate them. In the absence of a sensitivity analysis (see section 7 below) EPA objects to the use of the current values of ESC.

5. The Claim that Errors in ESC are Self-Correcting. In Section E.12.7.6.1, the Army presents an argument that errors or biases in ESC would be self-correcting, in the sense that they would not lead to systematic errors or biases in estimates of risk.

> "It is emphasized that the TC map comparison provides information on which BMF values provide the most realistic risk estimates under the current assumptions regarding exposure soil concentrations (ESC). That is, if the ESC model tends to overestimate ESC, then the BMF which implies the best prediction of TC will be an underestimate of the true BMF for the RMA. The focus of the TC map comparison is to evaluate the prediction of risk, i.e., the performance of BMF and ESC together, not the general validity of the BMF itself. The map comparisons are not intended to provide information as to the validity of the methods or the general applicability of the BMF values to other ESC models and other sites." (p. E-91, emphasis added).

We note in passing that this shows that the process of "predicting" tissue concentrations by multiplying ESC by estimates of BMF derived from ESC and observed tissue concentrations is circular. (See EPA dispute letter of June 16, 1993, for a discussion of this problem.)

Apart from issues related to calibration, EPA objects to this argument and to its implication that errors or biases in ESC would be self-correcting and hence unimportant in risk characterization. (i) The argument is inconsistent with the definition of ESC (see passage from Section C.1.4 quoted in the

Discussion section above). (ii) The argument depends on the unstated assumption that the relation between ESC and the true exposure concentration is the same everywhere. This is extremely unlikely to be true. The errors and biases in ESC, which EPA has pointed out and discussed in Sections 1-3 above, derive primarily from the wide spatial variability in soil contamination and the non-uniform sampling (aggravated by the Army's methods of calculation). These errors and biases are likely to be widely variable in time and space. Hence, the map comparison presented by the Army does not provide an unbiased (self-corrected) validation of the BMF values, as claimed in Section E.12.7.6.

<u>6. Treatment of Uncertainty in ESC</u>. Despite the central importance of ESC in generating estimates of BMF and consequent estimates of BCs and risks, the discussion of uncertainties in ESC in Appendix E is very brief. Section E.12.4.2.1 states that uncertainties in exposure ranges contribute to errors in ESC.

> "Samples were often taken from such areas of transition and this contributes substantially to the overall uncertainty in the ESC. Errors in estimating ESC are also expected to explain some of the lack of correlation observed" (p.E-75).

Section E.12.4.2.2 states that uncertainties remain after replacement of BCRL values.

> "Although an attempt was made to estimate the most likely concentrations for the BCRLs, a substantial amount of uncertainty is still present in the data sets after they have been processed by the BCRL interpolation method" (p. E.75).

Section E.12.4.2.3 states that in some cases interpolation of soil concentrations onto the grid resulted in uncertainty in ESC.

> "However, [biota tissue] samples were often collected in areas of rapid transition between high and low [soil] concentrations. ... It is emphasized that the SC interpolated estimates are highly uncertain in such areas of rapid transition in concentration and, therefore, the ESC estimates in these areas are also very uncertain" (p. E-77).

Apart from these brief statements, there is virtually no discussion of uncertainty in ESC within Appendix E. The quoted statements are entirely qualitative; there is no sensitivity analysis and no attempt to estimate the possible magnitude of uncertainty in ESC. There is no discussion of the propagation of errors in ESC into derived quantities such as BMF or BC, except for the claim on pp. E-90 and E-91 that errors would be self-correcting (see previous section of these comments). In the text of the IEA, Section 5.4 acknowledges that uncertainty in ESC "outweighs concerns about other model parameter uncertainties" (p. 5-10). However, ESC is not mentioned in the chapter on uncertainty (Section 6.3.2).

These concerns relate to other discussions in the IEA, including, but not limited to, the following sections:

Page ES-9, first full paragraph
Page ES-12, last paragraph, fifth and sixth sentences
Page 2-9, section 2.4.2, second paragraph
Page 5-10, second paragraph
Appendix C.1, Section C.1.9, page C.1-57, second paragraph
Appendix D, Section D.1.4.4

ISSUE #2: CHARACTERIZATION OF HAZARD INDICES (HIs) > 10

EPA reserved its right to invoke dispute resolution on this item in its "early dispute" letter of March 1, 1993, pending revisions to the IEA. The Army has responded to EPA's previous request to map all HIs>1 as well as HIs>10, however, the language used to characterize the relative risks posed by areas with HIs>1 versus areas with HIs>10 is not acceptable. There are numerous places in the text where the areas with an HI>1 are described as "upper bound" risk areas where the likelihood of an actual adverse effect is low, and areas with HIs>10 are characterized as "likely to represent true risk". The Army has provided no basis for this distinction. While areas with HIs>10 may pose a higher risk, areas with HIs>1 do pose a true risk to biota and this must be clearly acknowledged in the text.

In the IEA Chapter 4 (Ecological Risk Characterization), the Army includes maps depicting areas where the "Hazard Index" for terrestrial biota or for components thereof exceeds 1.0, 10, or 100 (Figures 4.7.1 through 4.7.16). Hazard Index (HI) is defined on page 4-2 as the sum of "Hazard Quotients" (HQ), where HQ_i, the hazard quotient for the i'th contaminant, is defined as the ratio of estimated exposure to reference exposure for the i'th contaminant. The Army directs interpretation of the maps in the following passage on page 4-23:

> "The maps depicting HIs greater than 1.0 represent <u>the upper bound of potential risks</u>, primarily on the basis of the <u>conservative assumptions</u> associated with the toxicological endpoints (MATCs and TRVs) and the interpolated estimated soil concentrations for BCRL samples. The maps depicting the HIs greater than 10 reflect <u>the most likely risks estimates</u> for RMA target receptors." (Emphases added)

The Army interprets in this passage that the procedures for estimating risk are overly conservative by one order of magnitude, primarily because of conservative assumptions in establishing MATCs and TRVs, and in estimating replacement values for soil BCRLs.

EPA disputes this interpretation, specifically the three phrases italicized in the quoted paragraph.

This dispute addresses, at a minimum, Sections 4.7.1.1, 4.7.2, 5.5.2, C.1.4.1, C.1.7, C.1.8.1, C.2.4, C.2.6, E.10 and E.11.3.

Basis for Dispute

1. HI > 1 represents upper bound of risks

We note initially that the statement quoted above from page 4-23 conflicts with statements elsewhere in the IEA about the degree of protectiveness of the HI. For example, page 1-44 states "[a]n HQ value greater than 1 indicates the presence of risk from a single chemical. An HI value greater than 1 indicates the potential for risk from the collective chemicals of interest, even though no single chemical may in and of itself pose a risk." EPA has been unable to find any discussion or analysis anywhere in the IEA that would support the claims that HI = 1.0 represents the upper bound of potential risks, or that HI = 10 reflects the most likely risk estimates. Likewise, we have been unable to find any discussion or analysis anywhere in the IEA that would support the claim that conservative assumptions associated with MATCs and TRVs and replacement values for BCRL values are sufficient to justify this statement.

2. "Conservative" assumptions in deriving MATCs and TRVs. The derivation of MATCs and TRVs is presented in Sections C.2.4 and C.2.6, respectively. Neither of these sections support the conclusion that MATCs and TRVs were derived using "conservative" assumptions. MATC is defined on p. C.2-36 as "the whole-body tissue concentration that is unlikely to be harmful to the individuals or populations of a species over prolonged exposure under field conditions." TRV is defined on p. C.2-50 in similar terms, except that the definition of TRV acknowledges uncertainty "spanning perhaps an order of magnitude or greater." The only elements in the derivation of MATC and TRV values that could be characterized as "conservative" are the Uncertainty Factors (UFs), whose derivation is presented in Section 2.4.2.5. There is nothing in this or other sections in the IEA that states or implies that these UFs are unreasonably conservative. Page C.2-43 states that "The UF values listed in Table C.2-13 were selected on the basis of consensus [among technical representatives of the OAS] after thorough discussion of the available literature and application of informed professional judgment. The total uncertainty values used result in MATCs and TRVs that are reasonably conservative relative to existing toxicological data" (emphasis added). EPA's technical representatives took part in these discussions and agree that the UFs do not incorporate any conservative assumptions beyond those implied in the definitions of MATC and TRV. EPA notes that the values used for UFs and their components (Table C.2-13) are much smaller than those used in human health risk evaluation, which are generally characterized as conservative. The discussion of uncertainty in Section E.10 noted that the assignment of UF

values based on professional judgment "may or may not reflect the true level of uncertainty" (p. E-56). This is reflected in the Army's conclusory statement in the text of the IEA (Section 5.5.2):

> Toxicity estimates used in the ERC, the MATC and TRV, are intended to be protective of biota populations, including bald eagles. They serve as the concentrations to which measured or predicted tissue or dose values, respectively, are compared to evaluate potential risk. Because of their importance, the uncertainty in both toxicity estimates was quantitatively incorporated into the final values used. These quantitative UFs addressed intertaxon variability, study duration, study endpoints, and other modifying factors, <u>but the degree to which they under- or overcompensated for these factors is uncertain</u>. Further, other sources of uncertainty (e.g., individual and species differences in responses to toxicity within the trophic box and differences between the test individuals and test conditions among studies and form those at RMA) were not quantified" (page 5-13, emphasis added).

This statement is inconsistent with the interpretation on p. 4-23 that conservative assumptions associated with MATCs and TRVs are sufficient to justify the statement that "HIs > 10 reflect the <u>most likely</u> risk estimates."

<u>3. "Conservative" assumptions in estimating replacement values for BCRLs</u>. We were not able to find a discussion or analysis in the IEA that would support the Army's claim that the interpolated estimated soil concentrations for BCRLs incorporate "conservative" assumptions, nor that such assumptions could lead to overestimation of biota risk by a factor of 10. Section C.1.4.1.1, which presents the methods used for deriving replacement values for BCRL soil concentrations, does not use the word "conservative" and does not state any assumptions. Section E.12.4.1.1, which discusses uncertainty arising from interpretation of BCRL data, mentions that the 1/2 BCRL method tends to have a positive bias when the proportion of BCRL samples is high. However, this method was used only for tissue samples and not for soil samples. This section discussed uncertainty in soil replacement values as follows:

> "The proportion of BCRLs for most of the trophic box-chemical combinations was high. In such cases the interpretation of the BCRLs through different statistical methods could have a significant impact on the fitted distribution; therefore, the uncertainty due to BCRLs was relatively high. The arithmetic estimators used to fit lognormal distributions were

chosen because they reduced the impact of uncertainty in BCRLs" (pp. E-69 to E-70).

Thus, the text discussion of this issue acknowledges relatively large uncertainty, but expresses the opinion that the impact of this uncertainty was reduced by the treatment of the data. This section does not mention conservative assumptions, nor does it state or imply that the uncertainties were more likely to reflect overestimation than underestimation.

Although the Army's own discussion of this issue in Section E.12.4.1.1 does not identify any conservative assumptions or biases, EPA believes that the Army's procedures result in upward bias (i.e., non-conservative) in the estimated values used to replace BCRLs. For a full discussion of this issue, see item 1 (Replacement of BCRLs) under "Calculation of ESC," above. This upward bias probably results in some overestimation of average soil concentrations in areas where BCRL replacement values contribute significantly to ESC -- including some of the areas outside the BSAs. This would result in some overestimation of exposures and risks if reliable estimates of BMF were used instead of those derived by the Army. However, there is no indication that the magnitude of this overestimation would be so large as to justify the Army's claim that HI = 10 would yield the "most likely risk estimates." Most of the criteria listed in Table 4.6-2 are much higher than the CRLs, so even replacing all BCRL values with the CRL could not erroneously yield values of HQ greater than 1. For the few cases where the criteria in Table 4.6-2 are lower than the CRL, an HQ value of 10 would not be an appropriate measure of risk.

In summary, EPA believes that the Army's claim that an HI of 10 would yield the "most likely" estimate of risk is inconsistent with the Army's own evaluation of risks and uncertainties, and inconsistent with the scientific evidence as evaluated in the IEA.

This comment also applies, at a minimum, to other portions of the IEA:
Page ES-11, top paragraph, last sentence
Page ES-12, last paragraph, second sentence
Page 4-25, first paragraph
Page 4-33, third paragraph
Page 6-16, Section 6.5.1

<u>Recommendations</u>

EPA recommends that the Army's characterization of the hazard index maps on page 4-23 be replaced by an accurate characterization of the risks posed to biota at various levels of HI and of the associated uncertainties. This characterization should include at least three points:

1. Based on best professional judgment, an HI of 1.0 represents the highest level of chronic exposure that is unlikely to result in adverse effects on populations exposed chronically in the field.

2. For values of HI greater than 1.0, the potential for adverse effects increases progressively, becoming large for HI > 10.

3. The range of uncertainty in these statements spans at least one order of magnitude. This uncertainty exists in both directions; hence, some risk may occur at values of HI as low as 0.1.

ISSUE #3: INTERPRETATION AND INCORPORATION OF OTHER BIOLOGICAL STUDIES

The IEA contains a number of unsubstantiated statements which state or imply that wildlife populations at RMA are "healthy" and/or unaffected by the current contamination. EPA believes that data from existing studies is insufficient to conclude whether RMA wildlife populations are affected by contamination or are "healthy." This dispute could be resolved by eliminating all conclusory statements concerning biota "health" from the IEA or providing substantiation for the statements.

Background

In our earlier comments on the Ecological Status and Health section (dated July 7, 1993), we stated that the Army drew conclusions from studies that were not designed to evaluate the assessment and/or measurement endpoints the Army had chosen and that the limitations of these studies should be clearly discussed in the text.

A review of the above-cited EPA comments prepared by USFWS on August 5, 1993, supports EPA's position and states:

> "Page 1, second paragraph. The last sentence states '(t)he cited studies can therefore be qualitatively discussed, but cannot quantitatively be used to infer that populations have not been impacted by contamination at RMA.' The Service agrees and would add that these studies cannot be used to infer that there are impacts to wildlife populations at the Arsenal either. With very few exceptions, these studies have not been designed to address the endpoints considered in this chapter and should therefore not be used for this purpose." (Emphasis added)

In its review of an earlier draft of the same chapter (letter from Donald R. Gober to William McKinney, February 18, 1993), USFWS states:

"The use of anecdotal information to make broad generalizations regarding wildlife health is inappropriate. Very little of the information referenced was gathered specifically to address wildlife health related questions. In order to make informed and rational decisions and statements regarding the health of aquatic and terrestrial wildlife populations on the Arsenal, studies must be designed and conducted with this specific intent."

and

"(T)he Service believes that in order to adequately assess contaminant effects upon fish and wildlife populations, community structure, or ecosystem health at the Arsenal, one must conduct comprehensive, ecologically integrated studies that are specifically designed to address contaminant effects upon fish and wildlife resources. To date, these studies have not been conducted on the Arsenal." (Emphasis added).

EPA concurs.

RMA is unusual among Superfund sites in that it contains extensive areas of little or no contamination as well as several extensive areas of high contamination.

In our July 7, 1993, comments on the Ecological Status and Health section we identified a concern about the areas in which wildlife studies had been performed. We indicated that many of these studies had been conducted outside of the more contaminated areas and were not designed specifically to evaluate species diversity or abundance in "contaminated" versus "uncontaminated" areas. Our concern was that this could bias conclusions drawn from the data toward depicting an abundant and diverse population in all areas of the Arsenal.

In its reply of August 4, 1993, USFWS stated:

"The statement is made that no attempt was made to evaluate species diversity and abundance in contaminated versus uncontaminated areas and, therefore, there may be biases drawn from the data. This is certainly true since the Service has made efforts to try to eliminate wildlife use from contaminated sites. This has included removing perching/roosting structures (such as large trees and power poles), planting of vegetative barriers, installation of visual barriers, distress calls and noise makers. Any survey done would in fact be biased by these management actions and; therefore, of little use." (Emphasis added)

Discussion

The Army has reviewed a number of biological studies conducted on RMA. These studies contain information of varying utility in an ecological risk assessment. The Army has summarized this information in Appendix C.5, and has incorporated its interpretation of many of these studies throughout the text of Volume I. EPA objects to the manner in which these data are currently used in the IEA/RC.

The IEA contains a number of vague and unsubstantiated statements which state or imply that wildlife populations and communities at RMA are "healthy" and/or unaffected by the current levels and distribution of contaminants, as summarized below:

> Ecological observations during the past decade indicate that the overall ecosystems and animal communities have retained their integrity and most wildlife populations appear healthy (p. 2-3).
>
> These adverse effects on individuals, however, are not apparent at the population level given the available data (p. 2-3).
>
> Any measure intended to mitigate risks at the trophic level will result in an overall improvement of the ecological integrity, even though the RMA ecosystem may not show any signs of dysfunction or reduced robustness currently because of its apparent high assimilative capacity for the present contamination (p. 4-7).
>
> Despite the effects of RMA contamination in individual animals, these effects are not apparent in the available data on wildlife populations at RMA (p. 4-31).
>
> ...the ecological measurement endpoints evaluated show that trophic diversity at RMA has not been adversely affected by chemical contamination. Studies on reproduction generally indicate a lack of adverse reproductive effects for birds and most individuals observed on RMA appear healthy (p. 4-32).
>
> Also, results of the ecological risk assessment indicate the ecological measurement endpoints evaluated relative to potential risks show that trophic diversity at RMA has not been adversely affected by chemical contamination and studies on reproduction generally indicate a lack of adverse reproductive effects for birds and most individuals observed on RMA appear healthy (p. 6-18).

Appendix C.5 on Ecological Status and Health contains a number of similar statements on individual species, as well as conclusory statements about populations, communities and ecosystems:

> Investigations on the effects of contamination at RMA during the past decade indicate that while some effects may still be present in biota at RMA, the overall ecosystems, animal communities, and wildlife populations appear healthy (p. C.5-57).

> Despite the effects of RMA contamination on individual animals, these [adverse effects of contamination] are not apparent in the available data on wildlife populations at RMA (p. C.5-57).

The Executive Summary includes a similar conclusory statement:

> The results of the ecological risk assessment indicate that potential risks occur in areas of RMA having elevated concentrations of contaminants; however, the ecological measurement endpoints evaluated show that wildlife diversity at RMA has not been adversely affected by chemical contamination. Studies on reproduction generally indicate a lack of adverse reproductive effects for birds and most individuals observed on RMA appear healthy (p. ES-13).

EPA considers all of these statements unsubstantiated. The available studies of wildlife at RMA are insufficient to draw <u>any</u> conclusions about ecological "health," or about populations, communities, or ecosystems. The lack of documented effects at these levels of ecological organization reflects lack of appropriate study, as indicated by USFWS, not lack of effects. Although many studies of wildlife have been carried out at RMA, as summarized in Appendix C.5, these studies were not designed to investigate ecological "health" or the integrity of populations and communities, and did not provide useful information on these subjects. EPA objects to statements in the text and in Appendix C.5 that state or imply otherwise without supporting documentation. Lacking pertinent studies, the ecological assessment must proceed using the documentation presently available.

Our concerns address Appendix C.5 and related statements incorporated throughout the document, particularly in Sections 2.1.3, 4.1.3, 4.9, 5.5.2 and 6.5. Additional points of concern are discussed below:

<u>1. Definition of Ecological "Health"</u>. Ecological "health" is defined in section C.5.1 as follows: "For purposes of this risk assessment, ecological health can be defined as consisting of the normal range of ecological characteristics that provide

the basis for selecting appropriate assessment endpoints (EPA 1989a)." This definition is inappropriate in the following ways:

(a) The citation is incorrect. The Risk Assessment Guidance for Superfund, Vol. II (EPA 1989a) did not define ecological health.

(b) The definition is so vague as to be meaningless. Assessment endpoints are defined on pp. C.5-9 to C.5-10 as "environmental characteristics, which, if they were found to be significantly affected, would indicate a need for remediation...". Thus the combined definitions assert that ecological health is the normal range of ecological characteristics that provide the basis for selecting appropriate environmental characteristics which (if affected) would indicate a need for remediation. This definition has no operational meaning except perhaps for the word "normal," which perhaps implies that a healthy ecosystem is one that is functioning "normally."

(c) The definition is stated to apply on a regional scale:

"In general, contaminant effects assessments apply to regional populations. These are populations that occupy a particular region characterized by a habitat or habitats that are more or less contiguous and occur within <u>a major biogeographic region</u> (e.g., shortgrass prairie and associated habitats such as riparian woodland, pastureland, and wetlands)." (p.C.5-1, emphasis added).

If this definition is accepted, then none of the information presented in Appendix C.5 has relevance. None of the studies was conducted on a regional scale or gave any information on regional populations on the scale specified. The definition given is inappropriate and should be replaced by a definition that specifies appropriate scales (see section 3 below) and specific characteristics that can be used to assess ecological endpoints (e.g., some of the characteristics listed on page C.5-10).

<u>2. Appropriateness of Off-Post Control Areas</u>. Pages C.5-1 and C.5-2 point out several characteristics of RMA which make it a "very unique site for ecological risk assessment." These include large size, proximity to a major urban area, extensive areas of native grassland, and sizable populations of deer, prairie dogs, and raptors. The uniqueness of RMA in this and other characteristics makes it very difficult to select appropriate off-post reference or control sites to serve as the basis for rigorous comparisons. Some of the studies cited in Appendix C.5 had no off-post control sites (e.g., great horned owls (p. C.5-30) and burrowing owls (p. C.5-32)). For most of the other studies, the locations of off-post control sites are stated, but inadequate information is given on their

characteristics or comparability with RMA. This information does not appear even in the original study reports, as cited in the IEA: small mammals (p. C.5-20), waterbirds (p. C.5-36), songbirds (p. C. 5-39), grasshoppers (p. C.5-40), earthworms (p. C.5-40). For aquatic snails (p. C.5-41), off-post control sites are referred to but are not identified or described. For fish (p. C.5-8), an off-post control lake is stated to have been similar to the study lakes in several characteristics, although not in management regime, but no specific information is given to support this statement. For American kestrels (p. C.5-22), several off-post reference and control sites were used, but are stated to have differed from the RMA in vegetation and habitat characteristics.

Thus, based on the information presented in Appendix C.5, rigorous comparisons with off-post control sites are not possible for any of the studies mentioned in the previous paragraph. Off-post control sites were either not used, not described in sufficient detail, or different from the study site. No conclusions (either positive or negative) about differences between RMA and control sites can be drawn from the information given in the IEA or in the original study reports. The conclusory statements based on the above-reference studies are unjustified.

<u>3. Spatial Scale and Within-RMA Comparisons</u>. As the IEA points out (p. C.5-1), RMA is unusual among Superfund sites in that it contains extensive areas with low contamination as well as several areas with high contamination. Thus, any conclusions about ecological "health" or about effects or lack of effects of contamination must be evaluated on a scale smaller than the entire site. Effects that may be occurring within or around the contaminated areas could be missed if studies and analysis are conducted on a larger scale. However, very few of the studies have been designed to investigate within-RMA differences in ecological characteristics. Most of the studies were conducted and reported on a RMA-wide basis: deer (p. C.5-16), bald eagles (p. C.5-26), great horned owls (p. C. 5.30), burrowing owls (p. C.5-32), raptors (p. C.5-33) and upland game birds (p. C.5-38). The study of songbirds was evaluated in a series of plots in different habitat divisions within RMA (p. C.5-39). The study of earthworms was limited to one contaminated site within RMA (p. C.5-40). In several of these studies, the distribution of animals was related to habitat features (deer, bald eagle, burrowing owl, raptors, songbirds). However, the studies were not designed to investigate differences related to contamination levels, and no systematic information on this relationship is presented or cited. Thus, no conclusions on within-site variations related to contamination can be drawn from any of the above studies.

Only four of the studies are stated in the IEA to have provided information on within-site variations that could be related to contamination. These studies are discussed in the next section.

4. Relationship of Ecological Characteristics to Contamination Levels. Most of the studies reported in Appendix C.5 were not designed to investigate the relationship between ecological characteristics and contamination levels, and did not provide systematic information on this relationship (see previous section). In several cases, conclusory statements are nevertheless made about this relationship. For example, p. C.5-20 states that "Average prairie dog density had no apparent correlation with degree of soil contamination." The source of this conclusion is the Biota RI, but the Biota RI only compared prairie dog densities in three areas classified according to likely contamination, without using any specific measurements or contamination levels. There was no comparison of the degree of contamination.

The section reporting on studies of American kestrels (pp. C.5-23 to C.5-26) contains several statements about apparent associations between soil contamination levels and distribution or reproductive success of kestrels. However, these associations are based only on geographical proximity, are not supported by any references to measurements of contamination levels or exposure, and apparently represent an after-the-fact pattern-seeking rather than testing of a priori hypotheses. Page C.5-23 states that in several years, nest outcome data were combined with egg and nestling samples for contaminant analyses, but no reference is made to any nest-by-nest analyses of these data or to attempt to correlate nest outcome with contaminant levels (p. C.5-25).

Of the three of the studies that provide systematic information on the relationships between patterns of soil contamination within RMA and differences in ecological characteristics the study of earthworms (p. C.5-40) involved collecting samples at an unspecified number of plots at one contaminated site and one on-post control site, and the study of terrestrial vegetation (p. C.5-43) is said to have detected major effects on vegetation in a few limited areas of high contamination, but no other obvious effects. The results of the earthworm study are cited as "not indicative of adverse contaminant effects." However, the use of only one contaminated and one control site means that the study had minimal statistical power to demonstrate effects that may have existed. The terrestrial plant study was reported too informally for evaluation.

Although information on contaminant levels in samples of many of the study species is given in the text of Appendix C.5

and in the tables in Attachment C.5.2, there is no attempt to identify patterns of contamination of the biota within RMA or to relate them to differences in ecological characteristics.

EPA concludes that the above-referenced studies, cited in Appendix C.5, do not provide conclusive information about the relationships between contamination and ecological characteristics. No conclusions (positive or negative) can be drawn about these relationships, based on the information given in Appendix C.5 or in the original study reports. The conclusory statements made in Appendix C.5 and throughout the remainder of the IEA are unjustified, and cannot be endorsed by EPA.

5. Population characteristics. Appendix C.5 includes a number of conclusory statements about "populations" of animals. The studies cited provide some information on population densities and reproductive rates, but little or no information about other population characteristics. The only species for which the total RMA breeding populations are stated are the mule deer (Biota RI) and the bald eagle (USFWS 1992). Arsenal-wide estimates of population density and area occupied are given for the black-tailed prairie dog (p. C.5-19). Breeding densities are given for four songbird species (Table C.5-3). Transect counts or other indices of population density are given for small mammals (p. C.5-21), raptors (p, C.5-34) and upland game birds (p. C. 5-38). Some numerical information apparently exists for other species, e.g., burrowing owl (p. C.5-32) and grasshoppers (p. C.5-40), but is not given in the IEA and was not located in the referenced citations. This information could be used as the basis for identification of temporal trends or spatial patterns, but the studies have not been designed to investigate variations related to contaminant distribution (see previous two sections of these comments).

Information on reproductive success is available for several species (deer, prairie dog, American kestrel, great horned owl, burrowing owl, waterbirds, and ring-necked pheasant). Except for the kestrel, however, there has been no attempt to identify within-RMA patterns of variation in reproductive success; for the kestrel, the relationship of these variations to contamination levels has been reported only very informally (see previous section of these comments). For all species, attempts to identify similarities or differences from off-post areas have not been systematic enough to draw reliable conclusions (see section 3 of these comments).

Population sizes are determined not only by reproductive rates, but by rates of mortality, immigration, and emigration. Mortality rates have not been measured in any study at RMA referenced in the IEA (a conclusory statement about "survival" of burrowing owls on p. C.5-33 appears to be erroneous). Immigration and emigration have not been investigated for any species in any study referenced in the IEA. Thus, there is insufficient documentation about any species to draw conclusions about population health, including viability at the population level.

6. Bias, Power, and Relevance. An unnumbered table on page C.5-46 classifies the various studies according to the Army's assessment of bias, power, and relevance. Most of the studies were classified by the Army as having low bias, medium power, and medium or high relevance. EPA disagrees with many of the Army's characterizations in this table and believes that the table systematically overstates the value of most of the studies for the purposes for which they are used. EPA's reasons are as follows:

(a) Bias. The Army's criteria for "Low" bias are given on p. C. 5-46 as "Samples representative [of sites of contamination]; controls [reference sites] were used and were appropriate." EPA does not agree that the samples were representative of sites of contamination for any study except probably that of grasshoppers or prairie dogs. As pointed out in item 4 of this comment, the relationship of the sampling sites to sites of contamination was apparently not part of the design of most of the other studies, soil contamination levels were not measured in relation to the animals that were investigated in the reported studies; the spatial relationship of the animals investigated to the known distribution of soil contamination was reported for only a few studies (e.g., kestrels and earthworms). Likewise, EPA does not agree that the control areas were "appropriate" for any of the studies. As pointed out in items 2 and 3 of this comment, control areas were frequently uncharacterized, inadequately described, or both. For these reasons, EPA believes that all the studies, as reported in Appendix C.5, had the potential for substantial biases (as well as for confounding with other factors). All should be classified as either "High" or "Medium" on the bias factor.

(b) Power. The Army's criteria for "High" power are given on p. C. 5-46 as "Study was designed to test contaminant-related effects; appropriate statistical tests were used; site data were compared to appropriate reference area data or regional background values." The criteria for "Medium" power are given as "Some combination of the above (professional judgment)." EPA disagrees with both of these criteria, for several reasons: (1) they partly duplicate the criteria for bias; (ii) they omit all the elements of statistical power (including sample size, sample variance, and assigned probabilities for rejecting the null and alternative hypotheses), (iii) regional "background" data are ill-defined and cannot be the basis for statistical tests. Because of the large variability in most ecological parameters, ecotoxicological studies generally have low power, unless they utilize very large sample sizes or measure endpoints with low variance. Based on these considerations, EPA would classify all the studies listed in this table as of "Low" power except for the studies of aquatic snails, grasshoppers, AChE inhibition, eggshell thinning, and kestrel reproductive success. Some or all of these could probably be classified as of "Medium" power, but the information given in the IEA is insufficient for classification in most cases except kestrels. The rabbit population study listed in the table is not mentioned in the text and may be an error. Other studies described in the text but not included in the table would all be classified as of "Low" power.

(c) Relevance. EPA generally agrees with the Army's classifications of "relevance" as given in the table. Some of the classification should be qualified to indicate that the studies are relevant only to some of the COCs. For example, eggshell thinning is relevant only to hypothesis about DDE, whereas several of the other endpoints measured have low relevance to DDE or mercury.

7. Bald Eagles. After discussing and summarizing studies of bald eagles, the Army presents the following conclusions:

> "The general health of bald eagles at RMA, low likelihood of their significant exposure, and their blood chemistry data do not suggest any adverse effects of RMA contamination." (p. C.5-29 to C.5-30).

This could be a correct conclusion, but it is not substantiated by the information presented in the IEA. Therefore, EPA has to consider this conclusion unsubstantiated, as explained below, regarding "general health", "low exposure", and "blood chemistry."

(i) The only information given about the "general health" of bald eagles is the statement that "[t]he majority of bald eagles captured at RMA have been within normal ranges for size, weight and condition for their age and the time of year they were captured." (p. C.5-29, emphasis added). No information is provided about the remaining "minority" of the eagles. Size, weight and condition are not good indicators of the major toxic effects of the COCs, which include impairment of reproduction and asymptomatic accumulation of contaminants in the body followed by lethal mobilization under conditions of stress. Page C.5-29 states that many of the eagles were trapped soon after arrival at RMA, i.e., before they would have time to accumulate the persistent COCs at the Arsenal. The information provided in the IEA therefore has little relevance for inferring any effects or lack of effects. No conclusion can be drawn.

(ii) The conclusion that there is "low likelihood of significant exposure" of the eagles is based on the "limited overlap between the areas of prairie dog habitation and the areas of highest contamination," the "somewhat limited" abundance of rabbits in the highly contaminated areas, the removal of prairie dogs and perch sites from Section 36 (Basin A), and the fact that bald eagles rely on theft of food items from other birds of prey (especially ferruginous hawks, which show a habitat-related avoidance of the most highly contaminated area) (pp. C.5-28 to C.5-29). Although prairie dogs may not occur in the highest contaminated areas, page C.5-20 stated that "[a]verage prairie dog density had no apparent correlation with the degree of soil contamination." Attachment C.5.2 lists high concentrations of dieldrin in many samples of prairie dogs and cottontails from several sections of RMA (not limited to Section 36).

Page C.5-36 and Attachment C.5.2 list high concentrations of dieldrin in several birds of prey, including lethal concentrations in the brains of single great horned owls, red-tailed hawks, and ferruginous hawks. Attachment C.5.2 (Figures 4-14 to 4-16) shows numerous sightings of raptors, including red-tailed hawks, rough-legged hawks, ferruginous hawks, and bald eagles, around and even within the more contaminated areas. The only eagle found dead on RMA and analyzed for COCs contained 0.276 ppm dieldrin and 1.70 ppm DDE in its muscle tissue, with concentrations in liver and brain also reported (Attachment C.5.2, Tables 5.1-4 and 5.1-7). These concentrations are similar to the whole-body MATCs of 0.41 and 2.17 ppm, respectively (Table C.2-14). Based on these data, one might conclude that bald eagles do have a significant likelihood of exposure; the length of residence at RMA could affect this conclusion.

(iii) Many of the blood samples were obtained soon after bald eagles arrived at RMA (p. C.5-29), so they provide little information about RMA exposure.

<u>8. Treatment of Uncertainty</u>. The treatment of uncertainty in the data and conclusions on ecological health and diversity is inadequate. In Appendix E, which presents the assessments of uncertainty for the entire IEA/ERC, uncertainty in the assessments of ecological health and diversity are not mentioned in any way, even though sections E.9 through E.13 are titled "Ecological Health." Likewise, in the main text (Volume 1) Section 5 (this section addresses uncertainty), does not mention uncertainty in the assessments of ecological health and diversity in any way. Section 6, (Integrated Endangerment Assessment), mentions ecological health only in the form of a conclusory statement ("trophic diversity has not been adversely affected..."p. 6-18). An almost identical statement appears in the Executive Summary (p. ES-13)

The only discussion of uncertainty and limitations in the characterization of the status and health of RMA fish and wildlife populations is given in three short paragraphs in Section C.5.5. The first two paragraphs list a number of sources of uncertainty without attempting to assess their significance. The third paragraph presents the Army's conclusions:

"Because of the variability in the design of past studies, the variability of wildlife exposure, and anticipated reduction in tissue concentrations, data from the long-term monitoring of population trends are needed. A consistent long-term study design would enable separation of contaminant effects on populations from natural long term population cycles. Such a study could minimize the number of variables that might obscure detection of any correlation between population trends or other ecological effects and contamination..." (pp, C.5-58 to C.5-59).

EPA agrees that more and better-designed studies would be required to separate contaminant effects on populations from natural long term population cycles, and to control for variables that obscure detection of correlations between population trends or other ecological effects and contamination. Despite the Army's acknowledgement in this paragraph of the need to achieve this separation and control, the remainder of the chapter fails to recognize that prior studies have not done so.

These concerns also relate to discussions throughout the remainder of the text, including, but not limited to, the following pages:

Page 1-3, section 1.1, fifth bullet
Page 2-9, section 2.4.2, last paragraph
Page 2-10, section 2.5, last paragraph, last sentence
Page 4-1, section 4.0, first paragraph, item 4
Page 4-6, last paragraph
Page 6-17, first full paragraph
Appendix C.5, in its entirety

Recommendations.

EPA considers that the information in Appendix C.5 is valuable in defining resources at risk and identifying where substantial populations exist near contaminated areas, after reflecting USFWS efforts to eliminate wildlife use from certain contaminated sites. Some of this information can provide the basis for monitoring of future trends, including changes that may take place post-remediation. However, the information cannot be used as per present statements in the IEA.

1. Appendix C.5 should be rewritten to include more accurate evaluation of the various studies and a more objective assessment of their limitations for the purposes of evaluating ecological "health" and diversity; conclusions one way or the other are not justified.

2. More actual data should be included in the Appendix in cases where the provision and assessment of data might justify better-defined conclusions (e.g., descriptions of control sites, statistical analysis of kestrel and small mammal data, blood

chemistry data on bald eagles).

3. A more appropriate definition of ecological "health" should be formulated and integrated into the assessment.

4. The table of assessments of bias, power, and relevance should be redrafted in light of EPA's comments.

5. The assessment of bald eagles should either be reevaluated, carefully qualified, or appropriate documentation provided in support of the conclusions currently stated.

6. Uncertainty in the results and conclusions should be discussed thoroughly and integrated into the conclusions.

7. All the conclusory statements quoted at the beginning of these comments (as well as conclusory comments on individual species in Appendix C.5) should be replaced by statements including appropriate qualifications and statements of limitations on each conclusion.

ISSUE #4: EXPOSURE ASSESSMENT FOR SHOREBIRDS

Shorebirds are a trophic box that raises special concerns because of the high concentrations of some of the COCs found in killdeers, the indicator species for this trophic box. Upper Derby Lake is currently maintained dry to prevent exposure to shorebirds (letter from USFWS re: Comments on Detailed Analysis of Alternatives, September 1, 1993).

In the Army's foodweb model, shorebirds are assumed to obtain food from both terrestrial and aquatic food chains. The IEA presents a complex procedure for partitioning exposure (represented either by calculated doses or by tissue concentrations) and risk between the terrestrial and aquatic components of the model. The presentation is confusing and possibly incorrect, and some of the results appear to be in error (please see below). The documentation is cryptic and EPA has been unable to reproduce the results in the limited review time available for the IEA. EPA reserves the right to extend this dispute to the other birds whose diet is partitioned (heron, waterbird, eagle) if the problem proves to be a model error rather than calculation or documentation specific to the shorebird.

This dispute primarily addresses Sections 4.6, 4.7, C.1.6, C.1.7, and D.1.3, and Figures C.3-13 and C.3-19.

<u>Basis for Dispute</u>

1. <u>Presentation problems</u>. The section titled "Risk Calculations for Trophic Boxes with both Terrestrial and Aquatic

Food Chains" (p. C.1-50 through C.1-56) correctly notes that the establishment of biota criteria (BCs) from data on animals in these trophic boxes is mathematically indeterminate unless the partitioning of exposure is specified independently. It proposes two methods for this partitioning (p. C.1-55). Example 1 is stated to be the method that "was done to calculate the BC used for the risk results given in Appendix C.3.2 ... This method is illustrated in Section D.1.3.2." However, there is no Appendix C.3.2, and Section D.1.3.2 is merely a list of spreadsheets; it does not illustrate the method. Example 2 is a description of a possible method for setting BCs based on remediation efficiency, but does not state whether or how it was used. Therefore, what was done is not clear and not possible to reproduce.

2. Results. Table 4.6-2 lists the soil criterion for aldrin/dieldrin for shorebird as 4.64 ug/g. This is one to two orders of magnitude greater than the soil criteria for aldrin/dieldrin for other trophic boxes. Table E.13.1 lists the PBC derived from BMF_{model} for aldrin/dieldrin in shorebirds as 0.0938. Different procedures used by EPA to verify the 4.64 criterion were not successful.

We believe that unless the value of 4.64 ug/g in Table 4.6.2 is in error, there would be an error in either the terrestrial/aquatic partitioning or in the calculations.

In Table C.2.10, ingestion of sediment is stated to comprise 16.0% of "diet." This is 60% of the aquatic component of the "diet" (10.5% food plus 16.0% sediment), an unrealistically high fraction. In contrast, ingestion of soil is not considered, although terrestrial food items comprise 73.5% of "diet." It is not clear whether and to what extent this apportionment in dietary fractions contribute to the results indicated above.

The value for endrin in Table 4.6-2 also appears to be in error.

Recommendations

This part of the model should be reviewed and the calculations repeated using correct input parameters. The presentation and documentation of the calculation should be rewritten to make clear what was done and to allow the reader to review and reproduce the procedures. If errors in procedure or calculations are found for the shorebird trophic box, the calculations for heron, waterbird and eagle should be reviewed also.

ISSUE #5: UNUSABLE SOFTWARE AND INCOMPLETE DOCUMENTATION

As supplied by the Army to EPA, the IEA/ERC software is not functional. A second implementation of the ERC model, using a computer spreadsheet and a program to generate frequency distributions that was used by the Army as part of the IEA analysis, was not supplied until September 10, which left only 10 calendar days for review. On September 10 we were informed that the software supplied with the IEA was no longer being used for biota calculations and this newly transmitted spreadsheet was in fact the basis of the ecological risk assessment. No previous notification of this fact had been given to EPA. A 20-calendar-day period is an extremely short amount of time to review a document as extensively modified as the IEA. The submission of non-functional software and the late submission of what remains an undocumented spreadsheet are not acceptable.

EPA's Office of Research and Development (ORD) has reviewed the BMF dispute issue. ORD's position paper dated August 18, 1993, states:

> "[o]ne of the basic tenets of science is that, to be supportable, enough information must be provided so that a reviewer can reproduce the findings. This is particularly important for highly complex or highly automated analyses as this."

Resolution of dispute on this subject will require an extended review period for the ecological calculations, for independent confirmation of the analyses.

ISSUE #6: CHARACTERIZATION OF UNCERTAINTIES RELATIVE TO BIOTA

Background.

In general, the characterization of uncertainties should be reassessed after dispute resolution is finalized. (See dispute item #1).

The Army presents the conclusion that overall uncertainty in ecological risk is about one order of magnitude and interprets that to be all one way: conservative. This position is not substantiated by the analysis presented.

The Army has provided a more comprehensive list of uncertainties in this IEA than in the previous draft. In many cases, however, there is little discussion regarding the impacts of these uncertainties, or their significance.

Uncertainties in the Ecological Risk Assessment are listed and discussed in a systematic manner in Appendix E. The conclusions are summarized briefly in Sections 5.4 and 5,5, and are tabulated in summary form in Table 5.0-1.

The analysis of uncertainty in these sections is insufficient, as described below. This dispute item primarily addresses Sections 5.4, 5.5, 6.5.1, and Appendix E (Sections E.9 through E.13).

In other sections of this review EPA presented its objections to the analysis of uncertainty in the calculation of ESC and the assessment of ecological health/diversity. Those are incorporated into this issue by reference. This issue does not address issues related to uncertainty in calibration, although other aspects of uncertainty in BMFs are addressed.

Basis for Dispute

1. General. The discussion of uncertainties relative to biota in Appendix E is almost entirely qualitative. Sections E.9 through E.11 list about 23 sources of uncertainty and characterize each in purely qualitative terms, devoting less than half a page of text to each factor (pp. E.52 through E.61). Table 5.0-1 (Volume 1) tabulates risk assessment components and classifies each according to its potential effect on the estimation of risk. Review of this table shows that of 37 risk assessment components applicable to biota, 22 components are classified as "Unknown", 6 as potentially overestimating risk, 5 as potentially underestimating risk, and 3 as having little impact. One component is classified as overestimating, underestimating, and "unknown." There is no attempt to quantify any of these sources of uncertainty, or the likely effect on the overall uncertainty in characterization of ecological risk.

In contrast, Section E.12 is an extensive (37 pages, plus many tables and maps) analysis of uncertainty in estimates of BMF. This section is fully quantitative and derives estimates of BMFs and Biota Criteria using three different methods. Section E.13 is a short summary purporting to present "Overall Uncertainty about Risk"; it concludes with Table E.13.1, which tabulates the three sets of estimates of BCs mentioned in the previous sentence. The conclusory statement about uncertainty in the text is in Section 6.5.1:

> "In interpreting the numerical results of the EA for RMA, it must be emphasized that the uncertainty considerations discussed above should be construed as a framework within which remediation decisions for RMA can be made. For example, based on the uncertainty evaluations performed, areas that differ by less than one order of magnitude in their level of risk may justifiably be considered to be essentially equivalent in prioritizing sites for remediation. In other words, care should be taken in defining remedial priorities among areas with HIs that vary only slightly from 1.0 (e.g., those with HIs less than 10)..." (p. 6-16, emphasis added).

The unbalanced treatment of uncertainty in these sections distorts the conclusions. The fact that only one factor is explored quantitatively, and that the final table presents the results of the assessment of this factor alone, yields the impression that only this factor is important. EPA does not agree that Table E.13.1 fully represents all the uncertainty resulting from methods of estimating BMF; it certainly does not represent the overall biota risk uncertainties resulting from all the other factors.

The conclusion (p. 6-16) that overall uncertainty is about one order of magnitude is not derived from Table E.13.1 and there is no other quantitative analysis on which it could be based.

2. Uncertainty in BMFs. Section E.12 discusses and analyzes uncertainties associated with biomagnification parameters. While parts of this discussion are useful, EPA views the overall treatment of uncertainties in BMFs as incomplete:

(i) EPA does not accept the Army's methods of calibrating the model, of "predicting" tissue concentrations, and of using maps of "predicted" tissue concentrations to evaluate the "performance" of different estimates of BMF. This is the subject of an ongoing dispute and will not be discussed further herein (see EPA's dispute letter of June 16, 1993, and ORD's position paper, transmitted September 1, 1993).

(ii) Although uncertainty in ESC distributions is acknowledged and discussed at some length in Section E.12.4.2, the contribution of this uncertainty to the overall uncertainty in BMF is not assessed anywhere in Section E.12. All the uncertainty analyses in this section are based on the same set of values of ESC. EPA's concerns with the calculation and use of ESC are presented in item 1.2 of these comments (Dispute issue No. 1).

(iii) Quantitative uncertainty analyses are presented in this Section for only four components of the calculation of BMF: the method used for replacing BCRL values (analyzed only for TC and not for ESC; Table 12-4); the assumed correlation between TC and ESC (Table 12-5); the screening of the data (analyzed only for its effect on the TC/ESC correlation; Tables E.12-2 and E.12-3); and the statistical method used (Table E. 12-6). Although the results of these uncertainty analyses are useful, many other components of the calculation contribute to the overall uncertainty in BMF. These components are treated qualitatively in Appendix E, resulting in understatement of the overall uncertainty in BMF.

(iv) Although literature-based BMF estimates are defined on page E-64, they (BMF lit) are not presented in the IEA. Hence, there is no opportunity to perform a "reality check" on the results of the statistical calculations.

3. Other Factors Inadequately Considered. Several other factors that contribute to uncertainty in the ERC are omitted from or inadequately considered in Appendix E:

(i) Target species. Section E.9.1 states that the target species were chosen "from species present that best represent the uptake...and contaminant transfer" of contaminants. While EPA agrees that the species chosen were reasonably representative of these processes, they do not necessarily include the species at greatest risk (for example, the food web model includes no predatory mammal such as the American badger). Uncertainty resulting from species selection is not adequately considered in Appendix E.

(ii) Use of home ranges. Sections E.9.3 and E.12.4.2.1 discuss uncertainty resulting from modeling of exposure ranges. These sections discuss variability and uncertainty in size, shape and location of the exposure range, but do not discuss variability in the way in which an animal uses its home range. Many animals focus their activities (especially feeding) in certain localities within the home range. This may be an important component of uncertainty in ESC.

(iii) Dietary fractions. Dietary fractions are fixed values in the calculation. Page E.98 acknowledges that uncertainty and variability in FR are not included in criteria distributions. Variability in diet may be an important factor, especially for some of the predatory species.

(iv) Sampling Program. Section 5.1.1 (Volume 1) discusses limitations associated with the RMA chemical database. This section addresses sampling design in two sentences only, indicating in one that soil borings were biased toward more contaminated areas, and in another, avoidance of agent/explosives (i.e., areas that presented special safety concerns). As discussed in numerous EA Subcommittee meetings and addressed by the technical panels, one of the single greatest uncertainties regarding the ecological assessment is that the sampling program was not designed to characterize biota exposure to contaminants. This is neither discussed in Section 5.1 nor presented in Appendix E. The discussion in Section E.9.3 minimizes the lack of exposure estimation, and the discussion in Section E.12.2 wrongly implies that the RMA data were co-located.

The uncertainty resulting from the lack of co-located soil and tissue samples within the actual exposure areas of target receptors (which has led to a large number of assumptions and complex statistical handling) must be assessed in the IEA.

4. Characterization of Overall Uncertainty. The only

OTHER COMMENTS

<u>Figures A.2-1 and A.2-1</u>: These figures present the site conceptual models for both human and ecological exposure. Several key exposure pathways are not represented. For example, these figures do not show that there is potential for direct exposure of either humans or biota to "Basins and Lagoons," "Buildings," or "Ordnance Testing and Disposal Areas." Additionally, the only exposure to soils is described as "spill sites, surficial soils and other near surface sites, and isolated contamination sites."

These two figures do not reflect conditions at RMA. Since these models are important tools for risk communication, they must be accurate. We recommend that they be revised and appropriately annotated.

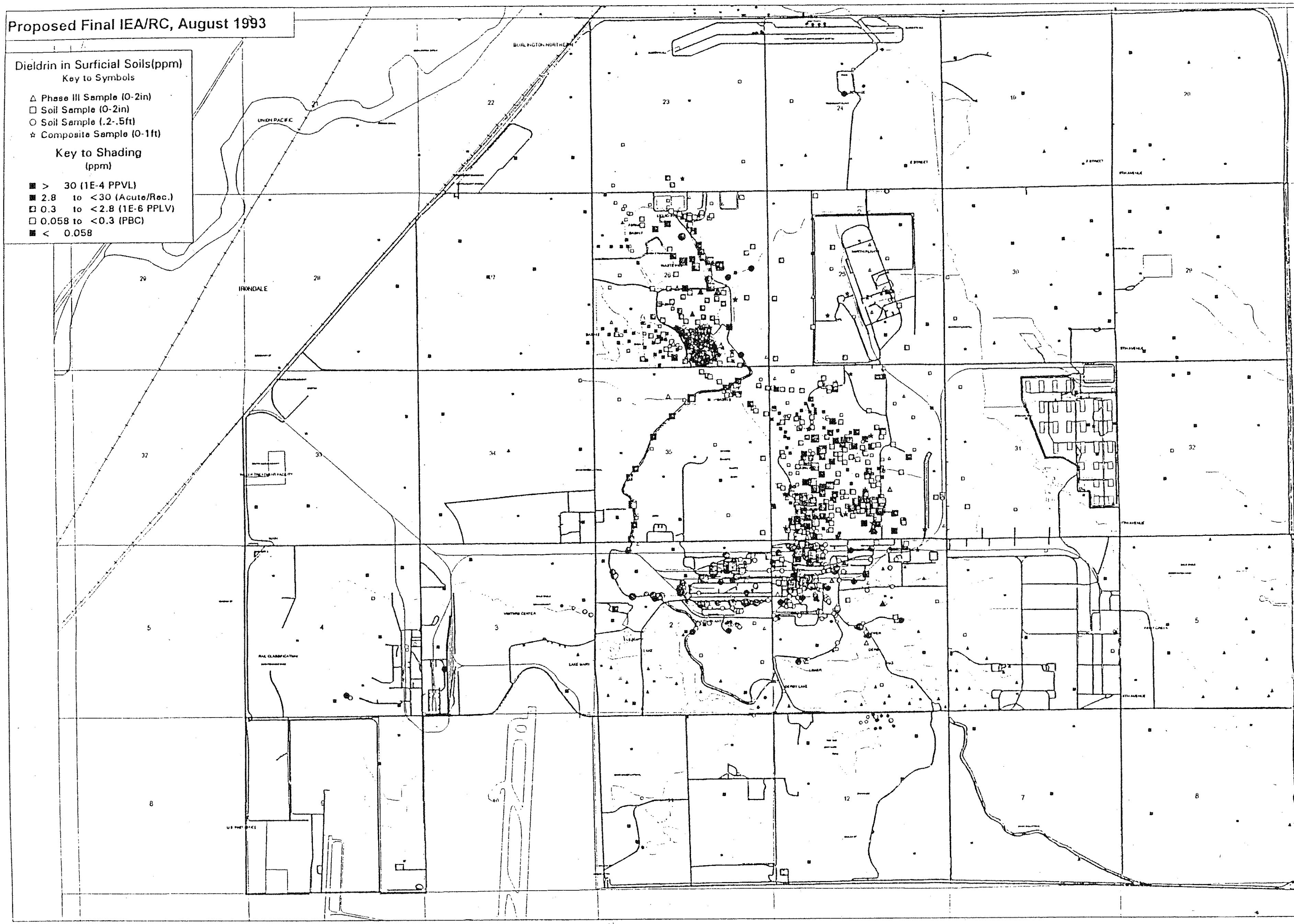
<u>Table B.1-10, Chronic noncarcinogenic dose-response data</u>: The RfD for lead is in error. EPA withdrew this RfD several years ago due to the unique dose response characteristics of lead.

<u>Table B.3-7, Age-specific daily soil ingestion rates</u>: Review of this table indicates that one of the data points (the child with high soil ingestion rates) was removed from the tabulation of the Stanek et al. data. We were not able to locate any rationale for removing this child's soil ingestion rate from the analysis. Please revise the table and associated calculations.

<u>Figure 4.7-9 and p. ES-12</u>: Figure 4-9 is said to illustrate soil HIs>1 based on inorganic concentrations greater than indicator levels. Although the Army did not subtract background values when calculating HIs, the Army did perform this subtraction to generate Figure 4-9. It is not appropriate to subtract background concentrations from locations where background was exceeded.

This figure shows inorganic concentrations that remain AFTER background concentrations are subtracted from the actual total concentrations. This is not appropriate because biota are exposed to the total concentrations, not the "total minus background" illustrated. A simple example explains this point. If a biota criterion was 15 ppm, a location with a total concentration of 20 ppm would have HI>1. However, if this location had a background of 10, it would not be shown as an exceedance. Figure 4.7-9 should be revised to include those locations where HI>1 based on total concentration, excluding only those locations where the total concentrations are at or below background.

Proposed Final IEA/RC, August 1993



characterizations of overall uncertainty are in Table E.13.1 and in Section 6.5.1. As pointed out in Section 1 of these comments, there is no apparent connection between these two characterizations, and the statement in Section 6.5.1. has no discernible basis. Table E.13.1 tabulates PBCs resulting from three alternative calculations of BMFobs. As the "bottom line" of Appendix E (and in fact, of the entire report), this is misleading. Presentation of the table in this location and in this manner, without any qualification, conveys that the Army believes that this table portrays the predominant or entire range of uncertainty in possible values of PBC. In fact, it takes into account only part of the range of one factor -- the method of calculation -- and does not represent uncertainty in any of the other parameters and procedures discussed in Appendix E.

This dispute issue also includes other discussions throughout the IEA/RC, including, but not limited to the following pages:

ES 12, last paragraph
ES-13, top paragraph
Page 2-11, first full paragraph
Page 4-4, Section 4.1.2.1, second paragraph, third sentence
Page 4-5, section 4.1.2.3, fifth sentence
Page 4-25, last paragraph
Page 4-26, top two paragraphs
Page 4-32, fourth paragraph
Section 5.0, in its entirety, especially the following pages:
Page 5-1, Section 5.1.1
Page 5-2, first full paragraph, last sentence
Page 5-5, second full paragraph
Page 5-9, first paragraph, second sentence
Page 5-10, second paragraph
Page 5-12, Section 5.5.1.2, second paragraph
Table 5.0-1
Page 6-12, first full paragraph
Page 6-13, second paragraph, last sentence
Page 6-13, Section 6.3.2
Page 6-15, first two bullets
Appendix D, Section D.1.4.4
Appendix E

ISSUE #7: UPDATED DESCRIPTION OF FORTUITOUS SAMPLES

In our dispute letter of March 1, 1993, we requested an updated presentation of fortuitous samples and measured contaminant concentrations. Such an updated presentation was not provided. The only material provided is on Tables C.5.1 (C.5.1-3 through C.5.1-24), which is a listing of samples collected under the Comprehensive Monitoring Program (CMP) only from 1988 to 1990. However, since 1990, we understand that dead or dying animals have been collected, for example: a badger with elevated dieldrin concentrations in 1992, a great horned owl in 1993 (see p. C.5-31), etc. The point of EPA's request was to obtain an evaluation of fortuitous samples updated to 1993, regardless of which program sponsored the analyses.

ISSUE #8: RESOLUTION OF DISPUTE CONCERNING ACUTE LEVELS FOR ALDRIN/DIELDRIN (HUMAN HEALTH)

The dispute concerning the acute reference dose for aldrin/ dieldrin (human health) was finalized on July 7, 1993. This enabled calculation of finalized acute criteria for these chemicals.

As indicated in our letters of December 7, 1992, May 20, and July 20, 1993, the Army was to recalculate and present the revised PPLVs for aldrin and dieldrin for all human exposure scenarios. We also recommended presentation in the IEA of a map outlining the areas at RMA that exceed acute levels for these chemicals, similar to the attached map, provided by CDH.

The main text of the IEA contained a table (Table 6.1-1b) giving the acute/subchronic values only for the industrial worker, but not for the other exposure scenarios, and did not contain the requested map. Appendix B.6 contained the levels for all scenarios.

The acute levels for only the worker scenario were presented in the main text. The acute levels for the other scenarios should be presented in close proximity, to enable comparison between scenarios. Likewise, the map showing RMA exceedances should be presented near the tabulated values.

ISSUE #9: POPULATIONS/ACTIVITIES PROTECTED

In our "early dispute" letter of March 1, 1993, EPA requested that the IEA include two sets of fixed input values which could provide the equivalent result, when used deterministically, as the 50 and 95 percentile probabilistic output.

These were not included in the IEA. On September 10, 1993, the Army's contractor supplied a table with the requested calculation for the biological worker/dieldrin. This table indicates that, when calculated as described above, the 50 percentile is achieved when all of the probabilistic input parameters (such as duration of exposure, soil ingestion, etc.) are set at 47% and the 95 percentile is achieved when all of the probabilistic input parameters are set at the 80th percentile. This is useful to conceptualize the PPLVs. A similar table, developed for each of the exposure scenarios, should be included in the IEA.